1

Object Detection in Images

Technical Field

The present invention relates to a method and system for detecting objects in
images, and particularly, but not exclusively, to a method and system which detects
moving objects taking into account lighting shadows and highlights of those objects.

Background to the Invention and Prior Art

The automatic detection of moving objects such as people or vehicles within
video images of a scene has been a goal of many researchers, and is a precursor to the
provision of automated classification or tracking applications. Additionally, automated
object detection systems are in themselves of use for monitoring and detection
applications. The actual step of discriminating between picture elements (pixels) within an
image which represent foreground or moving objects of interest and the background
scene is known generally within the art and herein as "segmentation".

One of the problems known in the art relating to the automated detection of
objects is of how to compensate for changes in lighting conditions, and in particular of how
to discern between an actual object, and a shadow or highlight that that object may cast or
otherwise contain. Especially within object classification or tracking systems, the need to
be reasonably certain that it is the object which has been detected and not its shadow is
important for subsequent matching steps, and hence techniques have been proposed in
the art which detect and remove segmented pixels caused by shadows and highlights.

More particularly, McKenna et al. in "Tracking Groups of People", *Computer
Vision and Image Understanding,* **80**, 42-56, 2000 describe a pixel segmentation
technique wherein an adaptive background image is employed, which recursively adapts
a background image to take into account changes in illumination (which are assumed to
be slow compared to object movement). A colour channel background subtraction
technique is then performed, wherein for any particular input image, the RGB channels of
the input image pixels are compared with the adaptive background, and dependent on the
results of a logical comparison of the respective input and background R, G, or B values a
pixel is set as either "foreground" or "background". The map of "foreground" pixels
constitutes a mask which is then used subsequently for further processing.

Due to it having been generated using a colour-difference background
subtraction technique the mask of foreground pixels contains pixels which represent
object shadows and/or highlights. Therefore, McKenna et al. also describe a second pixel

segmentation technique wherein shadows are detected using pixel gradient and chrominance information. More particularly, and as described in Horpraset et al. "A Statistical Approach for Real-time Robust Background Subtraction and Shadow Detection" *IEEE ICCV'99 FRAME_RATE workshop,* it is known that shadows exhibit a colour
5   constancy property in that the chromaticity of a pixel which is in shadow does not significantly differ from the chromaticity of that same pixel when it is not in shadow. Instead, the only change lies in the luminance of the pixel. This colour constancy is therefore used as a first discriminator by McKenna et al in that they assume that any pixel with a significant intensity change in comparison with the background, but without a
10 significant chromaticity change could have been caused by a shadow.

However, McKenna et al. also note that the above colour constancy discriminator fails when a foreground object is the same chromaticity as the background e.g. when black trousers cross a grey footpath. Therefore, McKenna et al. also describe using pixel gradient information of the input and background images to perform a comparison there
15 between, on the basis that if an input pixel is a shadow pixel, then its texture information should not have changed much from its background value. A pixel is thus flagged as foreground if either chromaticity or gradient information supports that classification.

Having performed pixel segmentation using the above technique, McKenna et al then perform a connected component analysis to identify connected objects. They note
20 that due to the chromaticity and gradient discrimination a connected object may contain "holes", in that pixels wholly located within the boundaries of the connected pixels and which should have been segmented as foreground are erroneously segmented as background, and hence not part of the connected object. To remedy this, reference is made to the "mask" produced by the RGB subtraction method described earlier, in that
25 each "background" pixel which is part of a hole in a connected object is set to foreground if the mask from the RGB subtraction indicates that it is foreground. Thus the holes within the connected objects may be removed, and foreground and background segmentation performed which takes into account shadows.

Whilst the colour constancy and gradient techniques described by McKenna et al.
30 are effective at identifying shadows for the purposes of a foreground-background pixel segmentation, the technique described by McKenna et al presents some problems when used in real situations due to the connected component analysis being performed before taking into account the "holes" due to the segmentation. For example, if the "holes" due to the segmentation actually overlap and completely bisect an object, then the connected
35 component analysis will not recognise the two blobs as a single object, but instead as two

3

separate objects. As a consequence, overlapping "holes" will not be recognised as such, and hence cannot be filled in by reference to the RGB subtraction mask. Conversely, until the connected component analysis is performed the "holes" themselves are not identified, and hence filled. The result of this paradox is that in some situations two smaller

5   connected objects may be identified where in reality only one object exists.

Summary of the Invention

        In a first aspect the present invention presents an image processing method for detecting objects within an input image, the image being composed of picture elements,

10 the method comprising:

        a) segmenting picture elements representing a foreground object within the input image from those picture elements forming the image background using a first segmentation technique, wherein the picture elements segmented as foreground include elements representing a shadow or highlight of the object; and

15   b) segmenting picture elements which have the characteristics of a shadow or highlight of an object from those picture elements representing the foreground object using at least one other segmentation technique adapted to detect shadows and/or highlights;

        the method being characterised by further comprising the steps of:-

20   c) segmenting as foreground surrounding picture elements to those picture elements which are already segmented as foreground;

        d) repeating step c) until picture elements which were not segmented as foreground after step a) would be or are segmented as foreground; and then

        e) detecting as objects groups of adjacent picture elements which have been

25 segmented as foreground.

        By performing the additional segmenting step c) before the detection of any objects using a connected component analysis or the like then any artefacts such as holes, bisections or the like in a group of adjacent segmented pixels representing an object which have arisen due to the shadow removal processing can be compensated for ·

30 by removal. In effect, the segmenting step c) is a morphological dilation operation which "grows" the skeleton blobs remaining after the shadow removal processing operation such that the blobs are merged back into one blob (in the case of a bisection), or any holes in the blob filled prior to the object detection step. The overall effect of the invention is thus to ensure that when the object detection step is performed then a single object will be

35 detected for each object in the scene, but that any shadows or highlights in the object will

4

not affect the object detection. The object detection technique of the invention is therefore robust to illumination changes causing shadows.

In a preferred embodiment of the invention the segmenting step a) preferably further comprises:

5   for each picture element of the input image which is to be segmented as foreground, validating the foreground segmentation by comparison of the picture element with probability models relating to surrounding picture elements, wherein the foreground segmentation is confirmed if at least half of the models indicate that the picture element is foreground. This helps to prevent "false foreground" segmentation of pixels, which may be

10   caused by noise in the input image, or by the image capture means such as a camera being unstable.

From a second aspect the present invention also provides a computer program or suite of programs arranged such that when executed on a computer system the program or suite of programs causes the computer system to perform the method of the first

15   aspect. Moreover, from a further aspect there is also provided a computer readable storage medium storing a computer program or suite of programs according to the third aspect. The computer readable storage medium may be any suitable data storage device or medium known in the art, such as, as a non-limiting example, any of a magnetic disk, DVD, solid state memory, optical disc, magneto-optical disc, or the like.

20   From a fourth aspect, the present invention also provides an image processing system for detecting objects within an input image, the image being composed of picture elements, the system comprising:-

image processing means arranged to receive an input image to be processed, and to apply the following image processing operations thereto:-

25   a) to segment picture elements representing a foreground or moving object within the input image from those picture elements forming the image background using a first segmentation technique adapted to detect differences in the input image from a general background image, wherein the picture elements segmented as foreground include elements representing a shadow or highlight of the object;

30   b) to segment picture elements which have the characteristics of a shadow or highlight of an object from those picture elements representing the foreground object using at least one other segmentation technique adapted to detect shadows and/or highlights;

and being characterised by further applying the following operations:-

5

c) to repeatedly segment as foreground surrounding picture elements to those picture elements already segmented as foreground until picture elements which were not segmented as foreground after step a) would be or are segmented as foreground ; and then

5   e) to detect as objects groups of adjacent picture elements which have been segmented as foreground.

Within the fourth aspect the same advantages, and same further features and advantages are obtained as previously described in respect of the first aspect.

Embodiments of the present invention improve upon the prior art discussed
10 above by providing an improved pixel segmentation technique which accounts for "holes" left in the segmentation due to the shadow processing prior to performing a connected component analysis to identify connected objects. This is achieved by applying a foreground segmentation technique as is known in the art to give a "mask" for the boundary of any segmented blobs (a "blob" being a group of adjacent segmented pixels),
15 and applying shadow processing techniques similar to those described by McKenna et al. to identify shadows and highlights in the input image. As also mentioned by McKenna et al. the shadow processing results in "skeleton" segmented blobs being left, which blobs may contain holes, be reduced in size, or even bisected in comparison to the blobs obtained from the background subtraction. Before applying connected component
20 analysis to the blobs to find connected objects a morphological dilation operation is applied to the skeleton blobs to dilate the pixels therein, the dilation operation being repeatedly applied so as to reconstruct the skeleton blobs until the re-constructed blobs touch the respective boundary of the mask of the corresponding blob (located at the substantially same position in the image as the blob or blobs being re-constructed)
25 obtained from the background subtraction. By "growing" the blobs using a morphological dilation operation then any holes or bisections in the "skeleton" blobs obtained from the shadow processing can generally be filled in to give a re-constructed blob, before connected component analysis is performed to identify the blob as an object. Moreover, by conditionally performing the dilation operation for any blob only until the respective blob
30 mask from the background subtraction step is touched or just overlapped, it can be assured that shadows will not be included in the re-constructed blob, because shadows are generally extended areas of the original object shape along a certain direction.

6

BRIEF DESCRIPTION OF THE DRAWINGS

Further features and advantages of the present invention will become apparent from the following description of an embodiment thereof, presented by way of example only, and by reference to the accompanying drawings, wherein:-

5   Figure 1 is a schematic block diagram of a system according to the embodiment of the present invention;

Figure 2 is a flow diagram illustrating the steps performed by the embodiment of the present invention;

Figure 3 is a vector diagram illustrating the principal of colour constancy used in the
10   embodiment of the present invention;

Figure 4 is a map of a morphological operator used in the embodiment of the present invention;

Figure 5 is an example adapted background image used in an embodiment of the present invention;

15   Figure 6 is an example input image with a shadow highlighted which may be used as input to the embodiment of the present invention;

Figure 7 is an example video image used as input to the embodiment of the present invention in an example;

Figure 8 is a segmentation mask image generated by a first segmentation technique
20   applied to Figure 7 within the embodiment;

Figure 9 is a segmentation map resulting from the removal of shadows and highlights by the embodiment; and

Figure 10 is a segmentation map corresponding to the image of Figure 7 after the morphological dilation operation of the invention has been applied.

25

DESCRIPTION OF AN EMBODIMENT OF THE INVENTION

An embodiment of the present invention will now be described with respect to the figures, and an example of the operation of the embodiment given.

Figure 1 illustrates an example system architecture which provides the
30   embodiment of the invention. More particularly, as the present invention generally relates to an image processing technique for detecting objects within input images, the invention is primarily embodied as software to be run on a computer. Therefore, the system architecture of the present invention comprises a general purpose computer 16, as is well known in the art. The computer 16 is provided with a display 20 on which output images
35   generated by the computer may be displayed to a user, and is further provided with

7

various user input devices 18, such as keyboards, mice, or the like. The general purpose computer 16 is also provided with a data storage medium 22 such as a hard disk, memory, optical disk, or the like, upon which is stored programs and data.

More specifically, on the data storage medium 22 are stored data 24
5   corresponding to an adaptive background image, data 28 corresponding to an input image, and data 30 corresponding to working data such as image data, segmentation maps, or the like used as intermediate storage during the operation of the invention. Additionally stored on the data storage medium 22 is executable program code in the form of programs such as the control program 31, a background adaptation program 32, a
10   foreground segmentation program 36, a connected component program 26, a shadow detection program 34, and a blob reconstruction program 38. The operation of each of these programs will be described in turn later.

The computer 16 is arranged to receive images from an image capture device 12, such as a camera or the like. The image capture device 12 may be connected directly to
15   the computer 16, or alternatively may be logically connected to the computer 16 via a network 14 such as the Internet. The image capture device 12 is arranged to provide sequential video images of a scene in which objects are to be detected, the video images being composed of picture elements (pixels) which take particular values so as to have particular luminance and chrominance characteristics. The colour model used for the
20   pixels output from the image capture device 12 may be any known in the art e.g. RGB, YUV, etc.

In operation, the general purpose computer 16 receives images from the image capture device 12 via the network, or directly, and then runs the various programs stored on the data storage medium 22 so as to process the received input image in order to
25   detect objects therein. A more detailed description of the operation of the embodiment will now be undertaken with respect to Figure 2.

Figure 2 illustrates a flow diagram of the operation of the embodiment of the invention. Starting at step 2.2, here the general purpose computer 16 receives an input image to be analysed using the invention from the image from the image capture device
30   12, either via the network 14, or directly. The embodiment of the invention then acts to analyse the contents of the received input image so as to detect foreground objects of interest therein. Within this embodiment, "foreground" objects are taken to be moving objects, or other objects of interest. The process of detecting foreground objects relies on effectively segmenting (i.e. distinguishing) the pixels representing the foreground objects
35   from pixels representing the background of the scene in the image, and then performing a

8

connected component analysis on the segmented foreground images in order to characterise them as objects.

The general purpose computer operates to process the received image under the control of the control program 31. That is, the control program 31 controls the order of

5 processing to be applied to received input images, and runs the other programs to perform their respective processing operations in turn, passing input data to them and receiving processed data from them as appropriate. In view of this, the first processing to be performed once the input image has been received is to perform a segmentation of those images that do not belong to the background scene.   Within the present

10 embodiment the adaptive background subtraction method proposed by Stauffer and Grimson in "Learning Patterns of Activity Using Real Time Tracking", *IEEE Trans on Pattern Analysis And Machine Intelligence, 22(8)*, August 2000 was adopted.   Here, a mixture of K Gaussian distributions is used to model the colour changes, at each pixel location, in the imaged scene over time.   With each incoming frame the Gaussian

15 distributions are updated, and then used to determine which pixels are most likely to result from a background process. This model allows a proper representation of the background scene undergoing slow lighting and scene changes, but not momentary variations such as swaying trees or flags in the wind.

In view of the above, within the present embodiment at step 2.4 foreground pixels

20 are segmented from the input image 28, using the Stauffer and Grimson technique, and at step 2.6 the adaptive background image 24 is updated to take into account the received input image, and the updated image stored in the data storage medium 22. Full details of the necessary procedures to undertake step 2.4 are described in Stauffer and Grimson, *ibid.* sections 3.1 and 3.2.   Regarding step 2.6, the background pixels are obtained as

25 follows: the pixel colours in the background image assume those of the incoming image if they are classified as background at step 2.4. In the case that the incoming pixels have been classified as foreground, then the mean of the Gaussian distribution with the largest weight at the lowest variance (the most probable background colour in the pixel) is chosen as the background pixel colour. As a result an updated background image can be

30 obtained for every incoming frame. Within the embodiment the segmentation step 2.4 is performed by the foreground segmentation program 36, and the background adaptation step 2.6 is performed by the background adaptation program 32, both under the control of the control program 28. An example background image from the sequence of input images used to generate the example images shown herein is shown in Figure 5. An

35 example input image with the shadow cast by an object circled is shown in Figure 6.

9

The result of the segmentation performed at step 2.4 is that a map of segmented pixels is obtained, as shown in Figure 8, for the example input image depicted in Figure 7. The foreground pixels thus obtained, however, are not exempt from false detections due to noise in the background or from a shaking camera. Hence, at step 2.8 a false

5   foreground pixel suppression procedure is introduced to alleviate this problem. Here, when a pixel is initially classified by the segmentation performed at step 2.6 as a foreground pixel, its eight connected neighbouring pixels models are examined. If the majority of these models, when applied to this pixel, agree that it is a background pixel, then it is considered as a false detection, and removed from the foreground segmentation.

10 This "false foreground" test is applied to every pixel which was segmented at step 2.4 using the Stauffer and Grimson technique, and hence the segmentation as foreground for each of these pixels is validated.

Following the "false foreground" validation, the map of segmented pixels as shown for example in Figure 8 is stored at step 2.10, for use as a "mask" of segmented

15 pixels later in the process. This map of segmented pixels is stored in the working data area 30 of the data storage medium 22.

Once the foreground pixels have been identified, a further detection technique is applied to find out if some of these foreground pixels correspond to regions likely to be cast shadows or highlights. As explained in Horprasert et al "A Statistical Approach for

20 Real Time Robust Background Subtraction and Shadow Detection", *ICCV'99 FRAME RATE workshop* at section 4.2, a pixel can be considered a shadow pixel if it has similar chromaticity to, but lower brightness than, those of the same pixel in the background image, or can be considered as a highlight pixel if it has similar chromaticity but higher brightness than the corresponding pixel in the background image. This characteristic of

25 shadow or highlight pixels is referred to as "colour constancy" in that the chromaticity of the shadow or highlight pixels in an input image is similar to or identical to the corresponding pixel in the background image, with the only difference being in the luminance of the input image pixels. Therefore, as a first step in detecting shadows and/or highlights at step 2.12 the colour distortion ("CD"), and brightness distortion ("BD")

30 is calculated as described by Horprasert et al and as shown in Figure 3 herein, between respective foreground pixels and the corresponding pixels of the adaptive background image stored at step 2.6, and possible shadows are detected. More particularly, at step 2.12 aggressive thresholds are used for the colour and brightness distortion, and in particular the following logical conditions are applied:-

35

10

```
IF CD < 10 THEN
    IF 0.5 < BD < 1.0   -> SHADOW
    IF 1.0 < BD < 1.25 -> HIGHLIGHT
ELSE FOREGROUND
```

5

In other embodiments of the invention the thresholds may differ from those given above; the precise values of the thresholds should be found by calibration of any particular implementation.

Those segmented pixels which are identified as either shadow or highlight pixels
10 by the above logical conditions are then marked for removal from the foreground segmentation map. Note that in the embodiment the above steps are performed by the shadow detection program 34, under the control of the control program 28.

Although the above colour constancy test is effective in detecting shadows, it was observed that the procedure was less effective in cases where the objects of interest have
15 similar colours to that of presumed shadows. To correct this, a second process comparing the gradients/textures of the regions of the segmented pixels is also used, in a similar manner to that described by McKenna et al at section 3.2 of their paper "Tracking Groups of People", referenced above. More particularly, at step 2.14 within the embodiment of the invention the shadow detection program 34 acts to calculate the
20 spatial gradients of those pixels of the input image which are marked as segmented according to the present segmentation map, and then compares the calculated gradients with correspondingly calculated gradients of the respective corresponding pixels from the adaptive background, as described in detail by McKenna et al in section 3.2 of their paper. The segmented pixels of the input image which are found to have a similar gradient to the
25 adaptive background are then marked as candidates for removal from the foreground segmentation map.

The result of the operations of the steps 2.12 and step 2.14 are that both shadow and highlight pixels are marked for removal from the segmentation map, but at this stage no segmented pixels have yet been removed from the segmentation map. At step 2.16,
30 therefore, the shadow detection program 34 identifies those pixels which were identified at both steps 2.12 and 2.14 as being candidates for shadow or highlight pixels, and removes those pixels from the segmentation map. As a consequence, those pixels which exhibited both colour and gradient similarity with the background are detected as shadows or highlights, and removed from the object segmentation.

11

As an aside, although within the preferred embodiment a pixel is only removed from the segmentation at step 2.16 if it was marked as a potential shadow/highlight pixel by both steps 2.12 and 2.14, in alternative embodiments this need not be the case, and a pixel may be removed from segmentation if either one or the other, or both, of the tests

5   performed at steps 2.12 and 2.14 indicates the pixel to be a shadow/highlight pixel. This may result in even more pixels being identified as shadow/highlight pixels and being removed from the foreground segmentation than in the preferred embodiment where both tests must be met.

Returning to the main embodiment, because of the aggressive application of the

10   shadow removal techniques (due to the thresholds chosen for the logical conditions) it will also be found that many segmented pixels which properly represent the objects of interest will also have been removed from the segmentation map. An example segmentation map produced by the application of steps 2.12 to 2.16 is shown in Figure 9 from where it will be seen, with comparison to the corresponding Figure 8 which illustrates the segmentation

15   map prior to the performance of the steps, that the segmented pixels representing the objects of interest are much reduced, and exhibit both holes and bisections within and of the group of segmented pixels representing the objects. However, it should also be noted that the segmented pixels which represent shadow pixels, which are very prominent in Figure 8, have been completely removed within Figure 9.

20   As will be seen from Figure 9, however, the map of segmented pixels stored at step 2.16 contains a speckle of erroneously segmented pixels therein, and hence at step 2.17 speckle noise removal on the segmentation map is performed in order to suppress isolated noisy foreground pixels. This speckle noise removal technique is performed using a morphological operator as shown in Equation 1:

25

$$\tilde{M} = M \cap (M \oplus N) \quad \text{Equation 1}$$

where N is the structuring element shown in Figure 4. From Figure 4, it will be seen that N is a 3 x 3 structuring element comprising the four-connected neighbours of the origin,

30   but that the origin is not included. Additionally, the term (M⊕N) in equation 1 identifies all the pixels that are four-connected to, i.e. next to, a pixel of M. Hence, $\tilde{M}$ identifies all the pixels that are in M and also have a four-connected neighbour and, therefore, these points correspond to all the points in M except the isolated ones. The resulting speckle free

segmentation map resulting from this speckle filtering step is then stored in the working data area 30 of the storage medium 22.

If a connected component analysis to find objects were to be performed on the blobs of Figure 9 following shadow removal, then due to the various holes and bisections
5   within the groups of pixels various erroneous objects would be detected. Therefore, the embodiment of the invention provides that the various holes and bisections in the segmented blobs of pixels are removed and this is performed at step 2.18 by the repeated application of a morphological dilation operation to the segmented pixels in the segmentation map, to segment surrounding pixels to the presently segmented pixels.
10 This morphological dilation operation is repeatedly performed until the segmented blobs touch or just overlap the boundary of the corresponding respective blob of segmented pixels produced from the segmentation at step 2.4 (i.e. Figure 8), whereupon the dilation operation is then stopped. The effect of this is to "grow" the segmented blobs left over after the shadow removal process so as to fill in any holes or bisections within the blobs of
15 segmented pixels.

More specifically, at step 2.18 a dilation operation using a 9 x 9 block structuring element (with each block of the element being set to "1") is performed on the foreground segmented pixels of the segmentation map corresponding to Figure 9 by the blob reconstruction program 38 to segment surrounding pixels to the already segmented pixels
20 as foreground pixels. Then, at step 2.20 the area of segmented blobs in the segmentation map thus obtained is compared with the segmented blobs within the stored mask, being the segmentation map obtained from the segmentation step of step 2.4 i.e. the Stauffer and Grimson segmentation. At step 2.22 an evaluation is performed to determine whether any of the segmented blobs obtained as a result of the morphological dilation touch or
25 overlap the segmented blobs of the mask segmentation map. If this evaluation determines that there is no touching or overlapping of the mask segmentation blobs, then processing returns in a loop to the morphological dilation step 2.18 which is performed again. The loop comprising the steps 2.18, 2.20, and 2.22 is repeatedly performed until the evaluation of step 2.22 returns positive, in which case the loop is then ended.
30 Formally, therefore, segmented blobs within the segmentation map are reconstructed according to Equation 2:

$$R = M_i \cap (\tilde{M} \oplus SE) \qquad\qquad \text{Equation 2}$$

where M_i is the mask image obtained from the segmentation of step 2.3, $\tilde{M}$ the segmented pixels after the shadow/highlight removal steps of s.2.12 to 2.16, and SE the

13

structuring element whose size usually depends on the size of the objects; as mentioned above a 9×9 square element proved to work well in all our tests, but other size structuring elements which result in a dilation may be used. The underlying idea of this morphological dilation step is that the shadow removed blobs keep at least a number of points that have

5   been robust to wrong shadow removal. These robust points are appropriate for leading the reconstruction of neighbouring points as long as they form part of the silhouette in the original blob obtained from the segmentation of step 2.4.

The resulting segmentation map output by the blob reconstruction program 38 once the processing loop has ended should resemble the example shown in Figure 10

10 from which it will be seen when compared to Figure 9 that the segmented pixel blobs have been reconstructed, and any holes or bisections therein have been removed. It will also be seen, moreover, that the reconstructed segmented pixel blobs of Figure 10 do not contain any substantial shadow pixels within the segmentation. The segmented pixel blobs thus obtained in Figure 10 can therefore be used as input to a classical eight

15 connectivity connected component analysis algorithm to link all the pixels presumably belonging to one object into a detected object, respectively. Thus, at step 2.24 a connected component analysis is performed on the segmentation map obtained from the processing loop by the connected component program 26, in order to detect objects from the segmented blobs contained within the segmentation map. Data identifying the

20 detected objects is then stored within the working data area 30 of the storage medium 22, and may be used subsequently as an input to an object classification application, an object tracking application, or the like. Alternatively, the embodiment of the invention may be used on its own solely to detect objects, and to display the objects on the display 20 to a user.

25 In other embodiments, a further action may be performed if one or more objects are detected in the input image, such as, for example, the computer 16 causing an output device to give an audio or visual indication of the detection of an object to a user. The embodiment of the invention is therefore particularly suitable for use in monitoring applications and the like.

30 Unless the context clearly requires otherwise, throughout the description and the claims, the words "comprise", "comprising" and the like are to be construed in an inclusive as opposed to an exclusive or exhaustive sense; that is to say, in the sense of "including, but not limited to".